

SYSTEM AND METHOD FOR PROVIDING CAPITALIZATION CORRECTION FOR UNSTRUCTURED EXCERPTS

Abstract

5 A system and method for providing capitalization correction for unstructured excerpts is described. An excerpt of unstructured content is tokenized into a set of words. The set of words is analyzed for correct capitalization. Individual characters constituting at least one such word in the set of words are evaluated. The at least one such word is skipped if determined to be of a predefined type.